2020

Distributionally Robust Max Flows

Chen, Louis L.; Ma, Will; Orlin, James B.; Simchi-Levi, David

SIAM

Chen, Louis L., et al. "Distributionally Robust Max Flows." Symposium on Simplicity in Algorithms. Society for Industrial and Applied Mathematics, 2020.
http://hdl.handle.net/10945/65899

Distributionally Robust Max Flows

Louis L Chen* Will Ma† James B Orlin‡ David Simchi-Levi§

Abstract

We study a distributionally robust max flow problem under the marginal distribution model, where the vector of arc capacities is random, with the marginals to the joint multivariate distribution being known, but the correlation being unknown. The goal is to compute the expected value of the max flow under the worst-case joint distribution of arc capacities. We provide a simple combinatorial proof that shows that for the case of finite-supported marginal distributions, this worst-case expectation can be efficiently computed, and moreover, the worst-case joint distribution can be explicitly constructed, despite being non-trivial in the sense that it is not a combination of monotonic or anti-monotonic couplings. Our technique is to use a related min-cost flow problem to generate a distribution over cuts in the graph, which in turn induces the worst-case joint distribution. It also provides an alternative interpretation of the problem as a zero-sum game between a capacity player and a cut player.

1 Introduction

Distributionally robust optimization is a vibrant research area concerned with computing the expected performance under the worst-case probability distribution from an uncertainty family. By allowing for nature to choose from a family of distributions, the decision-maker is "robust" against assuming a particular distribution which may be incorrect [6]. Surprisingly, such an approach also has optimization benefits, because for some structures of the uncertainty family, the expectation is easier to compute under the worst-case distribution.

We study one such structure in this paper, called the *marginal distribution model*. In this model, there are multiple unknown parameters. The marginal distributions for these parameters are given and assumed to be correct, but the correlation between them is unknown. The uncertainty family consists of all joint parameter distributions which are consistent with the given marginals. Such a formulation was first proposed in [11], and the desired expectation can be computed via strong duality results from convex analysis [7, 8, 4].

However, while the existing literature provides methods to find the objective value (the expected performance under the worst-case joint distribution) through a long sequence of duality results, methods to find the worst-case joint distribution itself are less developed. In this paper, we study a specific problem under the marginal distribution model, namely the *max flow* problem with unknown arc capacities. For this problem, we show how to explicitly construct the worst-case joint distribution. In the case of finite-supported marginals, this distribution has a polynomially-sized support. In fact, having N and A denote the set of nodes and arcs respectively, there is an accompanying distribution of cuts that can be described in $\mathcal{O}(|N|)$ space; equivalently, a distribution of cut-sets that can be described in $\mathcal{O}(|A|)$ space. In doing so, we also provide a simple, combinatorial proof of the aforementioned duality result for the max flow problem.

1.1 Literature Review Our model is related to that studied in works on the correlation gap and price of correlation [2, 1, 12]. Given a function acting on a random vector specified up to only its marginals, the correlation gap is the ratio between the expectation of the random function with respect to an extremal joint distribution (consistent with the marginals) and the expection of the random function with respect to the independent coupling of the marginals. Using assumptions on the function like monotonicity, submodularity, and the existence of a cost-sharing scheme, they are able to derive bounds on the correlation gap for all possible marginal inputs. By contrast, in our problem we focus on the exact computation of the expectation as well as the extremal distribution, instead of comparing against performance on the independent coupling. Furthermore, in our specific investigation into the value of the max-flow (equivalently, min-cut) as a function of the arc-capacities, our problem does not actually fall within the category of inputs that their bound pertains to.

The earliest work related to ours is [11], in which the

*Operations Research, Naval Postgraduate School, [Monterey, CA 93943], USA, louischenusa@gmail.com.

†Graduate School of Business, Columbia University, [New York, NY 10027], USA, wm2428@gsb.columbia.edu.

‡Sloan School of Management and Operations Research Center, MIT, [Cambridge, MA 02139], USA, jorlin@mit.edu.

§Department of Civil and Environmental Engineering, Institute for Data, Systems, and Society and Operations Research, MIT, [Cambridge, MA 02139], USA, dslevi@mit.edu.

problem of finding bounds on expected max flow, among other classical combinatorial optimization problems like shortest path, and network reliability is studied, and the marginal distribution model is introduced. Given marginal arc capacity distributions, [11] presents a convex program that computes the worst (over all joint distributions consistent with the marginals) expected max flow value, without knowledge of the worst-case joint distribution. But when it comes to characterization of a worst-case joint distribution, their study requires computation of an inefficiently-sized convex dual program. Indeed, merely formulating the constraints to the program requires identification of all s-t cuts; counting the number of s-t cuts alone is #P Hard [9].

The analysis presented in the subsequent study [7] closely mirrors our approach as well, in that they also utilize convex duality, but their problem of interest is in contrast the longest path problem, or the study of PERT network. They present a primal-dual pair of problems and utilize jointly a pair of primal and dual optimal solutions to compose a worst-case joint distribution. Unlike [11], their primal and dual problems are efficiently sized.

Our work presents a simple construction of the worst-case joint distribution for the setting of max flow. In similar fashion to [7], our construction will be based on the solution to an efficiently-sized convex dual program. In the case of finite-supported marginal arc capacity distributions, this program is efficiently solvable, providing an exact extremal joint distribution in polytime. Furthermore, in our primal-dual pair of problems, we recover the same convex program as in [11] that provides the computation of the worst-case expectation. Finally, our primal-dual formulation presents the intriguing interpretation of a two-person zero-sum game between a player deciding on arc capacities and a player deciding on an s-t cut. We establish the connection between this game's mixed Nash Equilibria and our distributionally robust max flow problem.

1.2 Outline of Problem and Solution

Let $G = (N, A)$ be a directed graph with a source and a sink. Each arc $(i,j) \in A$ has a random capacity $\tilde{u}_{ij}$ with a known marginal distribution F_{ij}. For any realization $\mathbf{u} = (\mathbf{u}_{ij})_{(i,j)\in A}$ of arc capacities, let $Z(\mathbf{u})$ denote the value of the max flow in G under arc capacities $\mathbf{u}$. Our *distributionally robust max flow* problem is to find a random vector $\tilde{\mathbf{u}}$ whose distribution F solves

$$\inf_{F\in\Gamma} \mathbb{E}_{\tilde{\mathbf{u}}\sim F}[Z(\tilde{\mathbf{u}})], \tag{1.1}$$

with Γ denoting the set of all joint (multivariate) distributions consistent with the given (univariate) marginals F_{ij} for each arc $(i,j) \in A$.

Pessimistic Bound. The first idea in our solution is to consider the following lower bound on (1.1). Suppose that instead of sending the max flow $Z(\tilde{\mathbf{u}})$ after seeing the realized capacities $\tilde{\mathbf{u}}$, we had to precommit to a flow vector (equivalently, a capacity vector) $\mathbf{w}$, and then whenever the realized capacity $\tilde{u}_{ij}$ was smaller than the flow w_{ij} we tried to send along an arc (i,j), we had to pay a penalty equal to the difference. Note that if $w_{ij} - \tilde{u}_{ij} > 0$ for two arcs along the same path in $\mathbf{w}$, then we still have to pay both penalties. Therefore, (1.1) is lower-bounded by the value of the following optimization problem:

$$\sup_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \Big\{ Z(\mathbf{w}) - \sum_{(i,j)\in A} \mathbb{E}_{\tilde{u}_{ij}\sim F_{ij}}[\max(w_{ij} - \tilde{u}_{ij}, 0)] \Big\}. \tag{1.2}$$

Re-interpreting Nature's Decision as a Distribution over Cuts. The second idea in our solution is to re-interpret nature's problem of choosing a joint capacity distribution F in (1.1) as choosing a distribution over cuts in the graph. Indeed, for any F, there is a corresponding distribution over (min) cuts whose cut capacities match the max flow $Z(\tilde{\mathbf{u}})$ on each realization of $\tilde{\mathbf{u}} \sim F$. Suppose that in this distribution, arc (i,j) crosses the cut with probability q_{ij}. If $\tilde{\mathbf{u}} \sim F$ is chosen optimally, then in each instance that arc (i,j) crosses the cut (equiv. (i,j) is in the cut-set), $\tilde{u}_{ij}$ must assume a value in its lower q_{ij}-th quantile. Otherwise, if not, then nature can re-couple F to decrease the expected cut capacity value while holding this distribution over cuts fixed. Therefore, nature's optimization problem is equivalent to finding a vector $\mathbf{q} = (q_{ij})_{(i,j)\in A}$ which solves

$$\inf_{\mathbf{q}\in\Delta} \sum_{(i,j)\in A} \int_0^{q_{ij}} F_{ij}^{-1}(p)dp, \tag{1.3}$$

with Δ denoting the set of all vectors q in $[0,1]^A$ consistent with some random cut-set $\tilde{Q} \subset A$ satisfying $Pr\,[(i,j) \in Q] = q_{ij} \quad \forall (i,j) \in A$

Solving Nature's Re-interpreted Problem using Min-cost Flow Duality. Finally, we show that nature's problem in (1.3) is exactly formulated and solved by the dual of problem (1.2), which can be rewritten as a min-cost flow problem. This leads to our main result.

THEOREM 1.1. *The optimization problems in* (1.1), (1.2), *and* (1.3) *all have the same value.*

Theorem 1.1 is proven in Section 3. An algorithm to solve the case of finite-supported marginals exactly and efficiently is presented in Section 4.1. We also interpret our main result using a game between a capacity player and a cut player, in Section 5.

2 Preliminaries and Notation

Let $G = (N, A)$ be a directed graph with node set N and arc set A. N consists of at least two nodes, where two are distinctly marked as the *source* s and the *sink* t. s has no arcs entering it and t has no arcs leaving it. As well, let there be given a nonnegative random variable $\tilde{u}_{ij}$ for each arc $(i, j) \in A$, with distribution function F_{ij}.

We will let $\mathbf{u} = (u_{ij} : (i, j) \in A)$ denote an arbitrary vector of arc capacities, and let $\tilde{\mathbf{u}}$ denote a random vector of arc capacities. For any $\mathbf{u} \in \mathbb{R}^A_{\geq 0}$, let $X(\mathbf{u})$ denote the *flow polyhedron* with capacities $\mathbf{u}$, defined as the collection of *flow vectors* $\mathbf{x} = (x_{ij} : (i, j) \in A)$ which satisfy

(Flow Balance Constraint)

$$\sum_{j:(i,j)\in A} x_{ij} = \sum_{j:(j,i)\in A} x_{ji}, \qquad \forall\, i \in N \setminus \{s, t\}$$

(Capacity Constraint)

$$0 \leq x_{ij} \leq u_{ij}, \qquad \forall\, (i, j) \in A.$$

That is, all nodes other than the source s and sink t must satisfy the constraint that the flow leaving it equals the flow entering it, and flows along all arcs are bound by their capacities. Let X denote the collection of flow vectors that satisfy the Flow Balance Constraints. Let $v(\mathbf{x}) = \sum_{j:(s,j)\in A} x_{sj} = \sum_{j:(j,t)\in A} x_{jt}$ denote the total flow sent from s to t by $\mathbf{x}$.

For any vector of arc capacities $\mathbf{u}$, let $Z(\mathbf{u})$ denote the *max flow* possible under capacities $\mathbf{u}$, defined as $\max_{\mathbf{x}\in X(\mathbf{u})} v(\mathbf{x})$. We are interested in computing a lower bound for $\mathbb{E}_{\tilde{\mathbf{u}}}[Z(\tilde{\mathbf{u}})]$, which holds over all joint distributions for $\tilde{\mathbf{u}}$ that are consistent with the given marginals $\{F_{ij} : (i, j) \in A\}$. Formally, a joint CDF $F(\mathbf{u})$, defined over $\mathbf{u} \in \mathbb{R}^A_{\geq 0}$, is consistent if it satisfies, for all $(i, j) \in A$ and $c \in \mathbb{R}$,

$$F((u_{ij} \leq c; u_{i',j'} \leq +\infty \;\; \forall\, (i', j') \neq (i, j))) = F_{ij}(c). \tag{2.4}$$

We will let Γ denote the set of all joint CDFs F satisfying (2.4). Then our problem, restated here, can be expressed as computing

$$\inf_{F\in\Gamma} \mathbb{E}_{\tilde{\mathbf{u}}\sim F}[Z(\tilde{\mathbf{u}})]. \tag{1.1}$$

We remark here that in the sequel, the use of "inf" and "sup" in the optimization expressions (1.1), (1.2), (1.3) will be replaced with "min" and "max", respectively. In the end, these replacements are fully justified and so for the sake of brevity we omit the technicalities for now, and we refer the interested reader to Section 5 where a formal explanation is presented.

We briefly recall and establish some probability notations here. Let the generalized inverse of the (i, j)-th marginal be given by $F^{-1}_{ij}(p) := \inf\{q : F_{ij}(q) > p\}$ for $p \in [0, 1)$, and let $F^{-1}_{ij}(1) := \inf\{q : F_{ij}(q) \geq 1\}$. Also, we will make use of the probability space $\Big((0, 1], \mathcal{B}, \lambda\Big)$, where $\mathcal{B}$ will denote the standard Borel sigma algebra on $(0, 1]$ and λ will denote the Lebesgue measure on $(0, 1]$.

We conclude this section by reviewing the fundamental duality relationship known as max-flow min-cut. Recall that an s-t cut, written $C = (S, T)$, is a partition of N into two subsets S (*source-side*) and $T = N \setminus S$ (*sink-side*) such that $s \in S$ and $t \in T$. The (s-t) *cut-set* of C is the set of arcs $\{(u, v) \in A : u \in S, v \in T\}$. Notationally, we will write $\mathcal{X}_{\text{cut}} := \{\chi \in \{0, 1\}^A : \chi \text{ is the characteristic vector to some s-t cut set}\}$, and if $Q \subset A$ is an s-t cut-set of some s-t cut, then $\chi_Q \in \mathcal{X}_{cut}$ will denote the characteristic vector of Q. Finally, $\Delta(\mathcal{X}_{\text{cut}})$ will denote the set of all (discrete) distributions H over $\mathcal{X}_{\text{cut}}$, and for any random vector $\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})$, we adopt the notation $q_{ij} := Pr(\tilde{\chi}_{ij} = 1)$, for any $(i, j) \in A$.

Now observe that the total flow sent from s to t, when the flow variables are set to $\mathbf{x}$, is given by $v(\mathbf{x}) = \mathbf{x}^\intercal\chi$ for any $\chi \in \mathcal{X}_{\text{cut}}$. In addition, for any $\chi \in \mathcal{X}_{\text{cut}}$, we let $\mathbf{u}^\intercal\chi$ denote the *cut capacity* for the s-t cut, or equivalently, the cut-set, corresponding to χ. As it turns out, these two concepts of total flow and cut capacity are tied by duality; informally, the classic Max-flow Min-cut duality states that the maximum amount of flow that can be sent from s to t is equivalent to the minimum cut capacity. This can be formalized:

Lemma 2.1. (Max-Flow Min-Cut) *Let $G = (N, A)$ be a digraph with source node s and sink node t in N. For each arc $(i, j) \in A$, let u_{ij} be a nonnegative arc capacity for arc (i, j).*

$$Z(\mathbf{u}) = \max_{x,v} v$$

$$\sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} = \begin{cases} v, & i = s \\ 0, & i \notin \{s, t\} \\ -v, & i = t \end{cases} \quad \forall i \in N$$

$$0 \leq x_{ij} \leq u_{ij}, \; (i, j) \in A$$

$$= \min_{\chi\in\mathcal{X}_{\text{cut}}} \sum_{(i,j)\in\mathcal{A}} u_{ij} \cdot \chi_{ij},$$

$$\textit{where } \chi_{ij} = \begin{cases} 1 & (i, j) \in \textit{cut-set} \\ 0 & (i, j) \notin \textit{cut-set} \end{cases}$$

This classical theorem immediately pays dividends insofar as allowing us to shed the optimization $Z(\tilde{\mathbf{u}})$ within the expectation for a term that appears tamer, albeit at the expense of a slightly more complicated decision: a joint random capacity and cut-set vector $(\tilde{\mathbf{u}}, \tilde{\chi})$.

Corollary 2.1.

$$\inf_{F\in\Gamma} \mathbb{E}_{\tilde{\mathbf{u}}\sim F}[Z(\tilde{\mathbf{u}})] = \inf_{(\tilde{\mathbf{u}},\tilde{\chi}):\tilde{\mathbf{u}}\sim F\in\Gamma, \tilde{\chi}\sim H\in\Delta(\mathcal{X}_{\text{cut}})} \mathbb{E}_{(\tilde{\mathbf{u}},\tilde{\chi})}\left[\tilde{u}^\intercal \tilde{\chi}\right] \tag{2.5}$$

In particular, if $(\tilde{\mathbf{u}}, \tilde{\chi})$ solves the right hand side, then $\tilde{\mathbf{u}}$ solves the left hand side, equivalently, (1.1).

3 A Simple Solution

We begin by examining a particular max-cost flow problem (w/ concave costs) that presents a lower bound [1].

Lemma 3.1.

$$\max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \left\{ Z(\mathbf{w}) - \sum_{(i,j)\in A} \mathbb{E}_{\tilde{u}_{ij}}[\max(w_{ij} - \tilde{u}_{ij}, 0)] \right\} \leq \min_{F\in\Gamma} \mathbb{E}_{\tilde{\mathbf{u}}\sim F}[Z(\tilde{\mathbf{u}})]$$

Proof. Since for all $w, \tilde{u} \geq 0$,

$$\underbrace{Z(w) - Z(\tilde{u})}_{\text{Difference in Max-Flow}} \leq \underbrace{\sum_{(i,j)\in A} \max(w_{ij} - \tilde{u}_{ij}, 0)}_{\tilde{u} \text{ to w expansion}},$$

the inequality follows. ∎

The max-cost flow problem in Lemma 3.1 can be rewritten as

$$\max_{(\mathbf{x},\mathbf{w}):\mathbf{x}\in X(\mathbf{w})} \left(v(\mathbf{x}) - \sum_{(i,j)\in A} \mathbb{E}_{\tilde{u}_{ij}}[\max\{w_{ij} - \tilde{u}_{ij}, 0\}] \right)$$
$$= \max_{\mathbf{x}\in X} \left(v(\mathbf{x}) - \sum_{(i,j)\in A} \mathbb{E}_{\tilde{u}_{ij}}[\max\{x_{ij} - \tilde{u}_{ij}, 0\}] \right),$$

which can be rewritten in the linear programming form (**Flow Problem**)

$$\max_{x\in\mathbb{R}^A, x_{ts}\in\mathbb{R}} \quad x_{ts} - \sum_{(i,j)\in A} \mathbb{E}\left[\max(x_{ij} - \tilde{u}_{ij}, 0)\right]$$

$$\sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} = \begin{cases} x_{ts}, & i = s \\ 0, & i \notin \{s,t\} \\ -x_{ts}, & i = t \end{cases} \quad \forall i \in N$$
$$x_{ij} \geq 0, \ (i,j) \in A,$$

[1] In fact, as we'll see in Theorem 3.1, this is exact.

wherein we have added to the original graph an arc (t,s) with arc cost of 1 and unbounded capacity.

In formulating the classical convex dual program (via "dualizing" the equality constraints), we arrive at the dual problem

$$\min_{\pi\in\mathbb{R}^N} \max_{x\in\mathbb{R}^A, x_{ts}\in\mathbb{R}} \Bigg(x_{ts} - \sum_{(i,j)\in A} \mathbb{E}\left[\max(x_{ij} - \tilde{u}_{ij}, 0)\right] + \sum_{i\in N} \pi_i \left(\sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} \right) + \pi_t x_{ts} - \pi_s x_{ts} \Bigg)$$
$$=$$
$$\min_{\pi\in\mathbb{R}^N} \Bigg(\max_{x_{ts}\in\mathbb{R}} (1 + \pi_t - \pi_s) x_{ts} + \sum_{(i,j)\in A} \max_{x_{ij}\in\mathbb{R}} \Big((\pi_i - \pi_j) x_{ij} - \mathbb{E}\left[\max(x_{ij} - \tilde{u}_{ij}, 0)\right] \Big) \Bigg)$$

We make the following observations. First, if $1 + \pi_t - \pi_s \neq 0$, then x_{ts} can be chosen such that the inner maximization is unbounded. Second, for any $(i,j) \in A$, the maximization problem can be solved explicitly in closed-form as

$$\max_{x_{ij}\in\mathbb{R}} \Big((\pi_i - \pi_j) x_{ij} - \mathbb{E}\left[\max(x_{ij} - \tilde{u}_{ij}, 0)\right] \Big) = \begin{cases} +\infty & \pi_i - \pi_j > 1, \text{ or } \pi_i - \pi_j < 0 \\ 0 & \pi_i - \pi_j = 0 \\ \int_0^{\pi_i - \pi_j} F_{ij}^{-1}(p)dp & 0 < \pi_i - \pi_j \leq 1. \end{cases}$$

The case of $0 < \pi_i - \pi_j \leq 1$ can be explained by the high-level intuition that at the optimal x_{ij}, the change in the second term $\mathbb{E}\left[\max(x_{ij} - \tilde{u}_{ij}, 0)\right]$ from incrementing x_{ij} should be equal to $\pi_i - \pi_j$; in other words, $F_{ij}(x_{ij}) = \pi_i - \pi_j$. To argue this formally, we can note that for any $(i,j) \in A$, the function $f_{ij}(x_{ij}) := \mathbb{E}\left[\max(x_{ij} - \tilde{u}_{ij}, 0)\right] \ \ \forall x_{ij}$ has $\left[\lim_{t\uparrow x_{ij}} F_{ij}(t), F_{ij}(x_{ij})\right]$ as its set of subgradients at x_{ij} (See [10]); hence, the optimality condition in the case of $0 < \pi_i - \pi_j \leq 1$ is

$$\pi_i - \pi_j \in \left[\lim_{t\uparrow x_{ij}} F_{ij}(t), F_{ij}(x_{ij}) \right],$$

so that it is optimal to set $x_{ij} := F_{ij}^{-1}(\pi_i - \pi_j)$, when $0 < \pi_i - \pi_j \leq 1$.

Hence, in the end, the dual problem (**Network Dual**) can be written:

$$\begin{aligned} \min_{\pi\in\mathbb{R}^N} & \sum_{(i,j)\in A:\pi_i > \pi_j} \int_0^{\pi_i - \pi_j} F_{ij}^{-1}(p)dp \\ & \pi_s - \pi_t = 1 \\ & \pi_i - \pi_j \leq 1 \ \ \forall (i,j) \in A, \end{aligned} \tag{3.6}$$

where the constraints are in place to avoid the aforementioned unbounded behavior.

We define $\Pi_{\text{feas}} := \{\pi \in \mathbb{R}^N : \pi_s = 1, \pi_t = 0, \pi_i - \pi_j \leq 1 \quad \forall (i,j) \in A\}$. Note that, without loss of generality, this is the feasible region to (**Network Dual**). And so the discussion above establishes that

$$\max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \left\{ Z(\mathbf{w}) - \sum_{(i,j)\in A} \mathbb{E}_{\tilde{u}_{ij}} \max(w_{ij} - \tilde{u}_{ij}, 0) \right\}$$

$$(3.7) \quad = \min_{\pi\in\Pi_{\text{feas}}} \sum_{(i,j)\in A:\pi_i>\pi_j} \int_0^{\pi_i-\pi_j} F_{ij}^{-1}(p)dp.$$

We will return to this relationship later from a game-theoretic perspective in section 5. But in the iterim, (3.7) suggests the following construction of a joint random vector $(\tilde{\mathbf{u}}, \tilde{\chi})$ that is feasible to (2.5). Fix an optimal solution π^* to (**Network Dual**) and consider the following random experiment procedure. Begin by drawing a random seed p uniformly from [0,1]. Then, each arc (i,j) is in the cut-set (i.e. $\tilde{\chi}_{ij} = 1$) if and only if $\pi_j^* \leq p \leq \pi_i^*$ (note that $\tilde{\chi}_{ij} = 0$ always if $\pi_j^* > \pi_i^*$). Finally, the capacity $\tilde{u}_{ij}$ of each arc (i,j) with $\tilde{\chi}_{ij} = 1$ takes on a value in its lower $(\pi_i^* - \pi_j^*)$'th quantile, i.e. equals $F_{ij}^{-1}(q)$ for some $0 \leq q \leq \pi_i^* - \pi_j^*$.

The correlated distribution over $\tilde{u}_{ij}$ (resp. $\tilde{\chi}_{ij}$) implied by the random seed $p \in [0,1]$ is in fact an optimal solution for problem (1.1) (resp. problem (1.3)) from Section 1.2. All of this is made precise in the theorem below, which is the formal version of Theorem 1.1.

THEOREM 3.1. *Let π^* solve (**Network Dual**), and let the random vector $(\tilde{\mathbf{u}}, \tilde{\chi})$ be defined on the probability space $\Big((0,1], \mathcal{B}, \lambda\Big)$ by*

$$\tilde{\chi}_{ij}(p) := \begin{cases} 1; & \pi_i^* > \pi_j^*, p \in [\pi_j^*, \pi_i^*] \\ 0; & \textit{otherwise}, \end{cases}$$

for all $(i,j) \in A$, and if $\pi_i^* \geq \pi_j^*$, define

$$\tilde{u}_{ij}(p) := \begin{cases} F_{ij}^{-1}(p - \pi_j^*); & \pi_j^* \leq p \leq \pi_i^* \\ F_{ij}^{-1}(\pi_i^* - \pi_j^* + p); & 0 < p < \pi_j^* \\ F_{ij}^{-1}(p); & \pi_i^* \leq p \leq 1, \end{cases}$$

otherwise if $\pi_i^* < \pi_j^*$ define $\tilde{u}_{ij}(\cdot) := F_{ij}^{-1}(\cdot)$.

Then $(\tilde{\mathbf{u}}, \tilde{\chi})$ solves (2.5) and hence $\tilde{\mathbf{u}}$ solves (1.1). It follows that

$$\max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \left\{ Z(\mathbf{w}) - \sum_{(i,j)\in A} \mathbb{E}_{\tilde{u}_{ij}} \max(w_{ij} - \tilde{u}_{ij}, 0) \right\}$$

$$(3.8) \quad = \min_{F\in\Gamma} \mathbb{E}_{\tilde{\mathbf{u}}\sim F}[Z(\tilde{\mathbf{u}})]$$

$$= \min_{\pi\in\Pi_{\text{feas}}} \sum_{(i,j)\in A:\pi_i>\pi_j} \int_0^{\pi_i-\pi_j} F_{ij}^{-1}(p)dp$$

Proof. First, we verify that the discrete distribution of $\tilde{\chi}$ is indeed a member of $\Delta(\mathcal{X}_{\text{cut}})$. Order the range of the π^* values as $0 = \pi^{(0)} < \pi^{(1)} < \ldots < \pi^{(M)} = 1$, where M is the number of distinct values. Then, for each $k \in \{0, 1, \ldots, M-1\}$, let $T_k := \{i \in N : \pi_i \leq \pi^{(k)}\}$, and $S_k := N \setminus T_k$. Consequently, each T_k contains node t and S_k contains node s, so that $C_k := (S_k, T_k)$ is an (s-t) cut with corresponding cut-set $A_k := \{(i,j) \in A : i \in S_k, j \in T_k\}$. It becomes clear then that $\tilde{\chi}$, as defined, is supported by this collection of cut-sets, with

$$\tilde{\chi} = \chi_{A_k} \quad w.p. \quad \pi^{(k+1)} - \pi^{(k)}, \;\; \forall k \in \{0, 1, \ldots, M-1\}.$$

Furthermore, as $T_0 \subset T_1 \subset \ldots \subset T_{M-1}$, and $S_0 \supset S_1 \supset \ldots \supset S_{M-1}$, the collection of (s-t) cuts C_k is sink-side (or source-side) nested.

Second, verifying that $\tilde{\mathbf{u}}$ has as distribution function one that is consistent with the given marginals, i.e., lies in Γ, is trivial.

Finally, we establish optimality of $(\tilde{\mathbf{u}}, \tilde{\chi})$. By the optimality of π^*,

$$\begin{aligned}
\mathbb{E}_{(\tilde{\mathbf{u}},\tilde{\chi})}\left[\tilde{\mathbf{u}}^\intercal\tilde{\chi}\right] &= \sum_{(i,j)\in A:\pi_i^*>\pi_j^*} \int_{\pi_j^*}^{\pi_i^*} F_{ij}^{-1}(p - \pi_j^*)dp \\
&= \sum_{(i,j)\in A:\pi_i^*>\pi_j^*} \int_0^{\pi_i^*-\pi_j^*} F_{ij}^{-1}(p)dp \\
&= \min_{\pi\in\Pi_{\text{feas}}} \sum_{(i,j)\in A:\pi_i>\pi_j} \int_0^{\pi_i-\pi_j} F_{ij}^{-1}(p)dp \\
&= \max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \left\{ Z(\mathbf{w}) - \sum_{(i,j)\in A} \mathbb{E}_{\tilde{u}_{ij}} \max(w_{ij} - \tilde{u}_{ij}, 0) \right\} \\
&\qquad\qquad \text{by (3.7)} \\
&\leq \min_{F\in\Gamma} \mathbb{E}_{\tilde{\mathbf{u}}\sim F}[Z(\tilde{\mathbf{u}})] \qquad \text{(Lemma 3.1)} \\
&= \min_{(\tilde{\mathbf{u}},\tilde{\chi}):\tilde{\mathbf{u}}\sim F\in\Gamma,\tilde{\chi}\sim H\in\Delta(\mathcal{X}_{\text{cut}})} \mathbb{E}_{(\tilde{u},\tilde{\chi})}\left[\tilde{u}^\intercal\tilde{\chi}\right], \\
&\qquad\qquad \text{(Corollary 2.1)}
\end{aligned}$$

as desired. ☐

3.1 Interpreting π^*: In the proof of Theorem 3.1, we establish that there exists a random cut $\tilde{C} = (\tilde{S}, \tilde{T})$ with corresponding random cut-set $\tilde{Q}$ such that $\tilde{\chi} = \chi_{\tilde{Q}}$, and the probability laws satisfy $Pr(i \in \tilde{S}) = \pi_i^* \; \forall i \in N$, $Pr(i \in \tilde{S}, j \in \tilde{T}) = Pr((i,j) \in Q) = \max(\pi_i^* - \pi_j^*, 0)$ $\forall (i,j) \in A$. With this, we see that for any node i, π_i^* represents the likelihood that node i lies in the source side $\tilde{S}$. Naturally, this is consistent with the fact that the source s has $\pi_s^* = 1$ and the sink t has $\pi_t^* = 0$, as required in the constraints in Π_{feas}.

We can now interpret $\int_0^{\pi_i^*-\pi_j^*} F_{ij}^{-1}(p)dp$ as yielding the expected contribution of the arc (i,j) to the random cut capacity value $\tilde{u}^\intercal \tilde{\chi}$ (under the extremal correlation). Furthermore, the totally-ordered relationship among the π^* variables in $[0, 1]$, means that $\tilde{C}$ has as support a collection of cuts that are sink-side (resp. source-side) nested, in that for any two potential realizations $\tilde{T}, \tilde{T}'$ (resp. $\tilde{S}, \tilde{S}'$), either $\tilde{T} \subseteq \tilde{T}'$ or $\tilde{T} \supseteq \tilde{T}'$ (resp. $\tilde{S} \subseteq \tilde{S}'$ or $\tilde{S} \supseteq \tilde{S}'$).

4 Algorithmic Procedure for $\tilde{\mathbf{u}}$

While the discussion immediately preceding Theorem 3.1 all but spells out an algorithm for the reader by describing a random experiment defining $(\tilde{\mathbf{u}}, \tilde{\chi})$, we formalize the design of an algorithm here for completeness sake. We present an efficient algorithmic procedure (based on Theorem 3.1) to find an exact solution (support along with probability mass values) to (1.1) for arbitrary s-t networks with capacity marginal distributions that are finite-supported. This same procedure can be adapted to create an efficient method for finding (using careful discretization) an approximate solution to (1.1) in the case of arbitrary marginal distributions.

4.1 General s-t Network and Finite-Supported Marginals In this section, we focus on the case of finite-supported marginals. More formally, we let arc capacities be 0 or take one of K possible positive values, sorted $0 = c_0 < c_1 < \ldots < c_K$. Each arc (i, j) has a randomly-initialized capacity given by a random variable $\tilde{u}_{ij}$ which has CDF F_{ij}, defined as $F_{ij}(c_k) = \Pr[\tilde{u}_{ij} \leq c_k]$ for all $k = 0, \ldots, K$ ($F_{ij}(c_K)$ is always 1), and pmf $f_{ij} = Pr[\tilde{u}_{ij} = c_k]$ for all $k = 0, \ldots, K$.

In this case, we can formulate the dual (**Network Dual**) to (**Flow Problem**) as a linear program in the following way. Noting that (**Flow Problem**) in this case is a max-cost flow problem with piecewise-concave costs, we can perform the standard transformation [3] to a capacitated max-cost flow problem with linear arc costs by replacing each arc $(i, j) \in A$ with K parallel arcs. More precisely, for any $(i, j) \in A$, in place of arc (i,j) we now have K parallel arcs directed from i to j, where for any $k \in [K] := \{1, 2, \ldots, K\}$, the k-th parallel arc has profit equal to $-\sum_{\tau=1}^{k-1} f_{ij}(c_\tau) = -F_{ij}(c_{k-1})$ and upper capacity equal to $c_k - c_{k-1}$. In addition, for the sake of analysis, we add the following arcs wherein it will never be profitable to have nonzero flow:

- For any node $i \notin \{s, t\}$, add an arc from i pointed to t, with cost = -1, capacity = $+\infty$
- For any node $i \notin \{s, t\}$, add an arc from t pointed to i, with cost = 0, capacity = $+\infty$

The linear programming dual (**Network Dual'**) is

$$\begin{aligned} \min_{\pi,\lambda} \quad & \sum_{(i,j)\in A} \sum_{k=1}^{K} (c_k - c_{k-1}) \cdot \lambda_{ij}^k \\ & 1 + \pi_t - \pi_s = 0 \\ & -\sum_{\tau=1}^{k-1} f_{ij}(c_\tau) + \pi_i - \pi_j \leq \lambda_{ij}^k \quad \forall (i,j) \in A, k \in [K] \\ & \pi_t - \pi_i \geq -1; \quad \forall i \in N \setminus \{s,t\} \\ & \pi_i - \pi_t \geq 0; \quad \forall i \in N \setminus \{s,t\} \\ & \lambda_{ij}^k \geq 0; \quad \forall (i,j) \in A, k \in [K] \\ & \pi_i \text{ free} \quad \forall i \in N \end{aligned}$$

Without loss of generality, we may impose the constraint that $\pi_t = 0$ and $\pi_s = 1$. A solution (π^*, λ^*) will yield a π^* that can be used to construct a solution as outlined in Theorem 3.1.

We now present Algorithm 4.1 which takes as input an s-t network $G = (N, A)$, along with marginal arc capacity distributions $\{F_{ij}\}_{(i,j)\in A}$ supported on $\{c_0, \ldots, c_K\}$. The algorithm efficiently finds and stores both the support and probability mass values for a discrete random vector $(\tilde{\mathbf{u}}, \tilde{\chi})$ solving (2.5), making use of the optimal solution to (**Network Dual'**). More precisely, it outputs a list of 0/1 arc-incidence vectors $X \subset \mathcal{X}_{\text{cut}}$, along with a corresponding list of probability masses $P_X \subseteq [0, 1]^{|X|}$. As well, it outputs a list of joint realizations of the arc capacities $U \subset \{c_0, c_1, \ldots, c_K\}^A$, along with a corresponding list of probability masses $P_U \subseteq [0, 1]^{|U|}$.

ALGORITHM 4.1.

$X \leftarrow \{\}$, $P_X \leftarrow \{\}$, $U \leftarrow \{\}$, $P_U \leftarrow \{\}$
$\pi^* \leftarrow$ Solve (**Network Dual'**)
$\Pi \leftarrow \text{SORT}(\{\pi_i^* : i \in N\})$
$P \leftarrow \text{SORT}(\{F_{ij}(c_k) : (i,j) \in A, k = 1, \ldots, K\} \cup \{\pi_i^* : i \in N\})$
for $k = 1 : \Pi.length() - 1$ **do**
  $T \leftarrow \{i \in N : \pi_i^* \leq \Pi[k]\}$, $S \leftarrow N \setminus T$
  $x \leftarrow \text{zeros}(|A|)$
  for $(i,j) \in A$ **do**
    if $i \in S$ and $j \in T$ **then**
      $x[(i,j)] \leftarrow 1$
    end if
  end for
  $X.\text{insert}(x)$, $P_X.\text{insert}(\Pi[k+1] - \Pi[k])$
end for
for $p = 1 : P.length() - 1$ **do**
  $u \leftarrow \text{zeros}(|A|)$
  for $(i,j) \in A$ **do**
    if $\pi_i^* \leq \pi_j^*$ **then**
      $u[(i,j)] \leftarrow F_{ij}^{-1}(P[p])$

else
  if $P[p] \in [\pi^*_j, \pi^*_i]$ **then**
    $u[(i,j)] \leftarrow F^{-1}_{ij}(P[p] - \pi^*_j)$
  end if
  if $P[p] \in (0, \pi^*_j)$ **then**
    $u[(i,j)] \leftarrow F^{-1}_{ij}(P[p] + \pi^*_i - \pi^*_j)$
  end if
  if $P[p] \in [\pi^*_i, 1]$ **then**
    $u[(i,j)] \leftarrow F^{-1}_{ij}(P[p])$
  end if
end if
end for
U.insert(u), P_U.insert($P[p+1] - P[p]$)
end for

See Figure 2 for an illustration of the algorithm's construction of $\tilde{\chi}$. Note that the corresponding joint arc capacity distribution $\tilde{u}$ is non-trivial in that it is not a combination of monotonic and anti-monotonic couplings. Indeed, arc $(s,2)$ appears in the cut-set with probability 2/3, and when it does (taking on a value in its bottom $\frac{2}{3}$ percentile), either arc $(s,1)$ must take on a value in its bottom $\frac{1}{4}$ percentile, or arc $(1,2)$ must take on a value in its bottom $\frac{5}{12}$ percentile, but *not both*.

5 Capacity versus Cut: A Two-Person, Zero-Sum Game

In this section, we outline a non-cooperative game that is integrally connected to the problem of (1.1), in order that we may provide some contextual meaning to the pieces of equation (3.8).

We consider two opposing players: the *capacity* player and the *cut* player. Given any directed graph $G = (N, A)$ and random variable collection $\{\tilde{u}_{ij}\}_{(i,j)\in A}$ as above, the capacity player chooses a nonnegative vector of arc capacities, and the cut player chooses an s-t cut-set. More formally, the capacity player's strategy space is $\mathbb{R}^A_{\geq 0}$, while the cut player's strategy space is $\mathcal{X}_{\text{cut}}$. The capacity player's utility $U(\cdot,\cdot)$ as a function of $\mathbb{R}^A_{\geq 0} \times \mathcal{X}_{\text{cut}}$ is given by

$$U(\mathbf{w}, \chi) := \mathbf{w}^{\intercal}\chi - \sum_{(i,j)\in A} \mathbb{E}\left[\max(w_{ij} - \tilde{u}_{ij}, 0)\right],$$

with the cut player's utility given by $-U$. Put another way, when the capacity player chooses $\mathbf{w}$ and the cut player chooses χ, the cut player pays the capacity player the cut capacity $\mathbf{w}^{\intercal}\chi$. In return, the capacity player, for each arc $(i,j) \in A$, pays the cut player the average amount that the capacity decision w_{ij} exceeded the random capacity vector $\tilde{u}_{ij}$.

If the capacity player chooses first followed by the cut player, the resulting utility is:

$$\max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \min_{\chi\in\mathcal{X}_{\text{cut}}} U(\mathbf{w}, \chi).$$

If, on the other hand, the cut player chooses first followed by the capacity player, the resulting utility is:

$$\min_{\chi\in\mathcal{X}_{\text{cut}}} \max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} U(\mathbf{w}, \chi)$$

That these two quantities may not be equal, tells us that this game does not necessarily contain a pure Nash Equilibrium. Indeed, consider the example in Figure 3. It can be verified that for this input,

$$\begin{aligned}
&\max_{\mathbf{w}\geq 0} \min_{\chi\in\mathcal{X}_{\text{cut}}} U(\mathbf{w}, \chi) \\
&= \max_{w_{s1}, w_{12}, w_{2t}\in[1,2]} \Big(\min(w_{s1}, w_{12}, w_{2t}) - 2/3(w_{s1} - 1) \\
&\qquad - 2/3(w_{12} - 1) - 2/3(w_{2t} - 1)\Big) = 1.
\end{aligned}$$

On the other hand,

$$\min_{\chi\in\mathcal{X}_{\text{cut}}} \max_{\mathbf{w}\in[c_0, c_K]^A} U(\mathbf{w}, \chi) = 2 - 2/3 = 4/3.$$

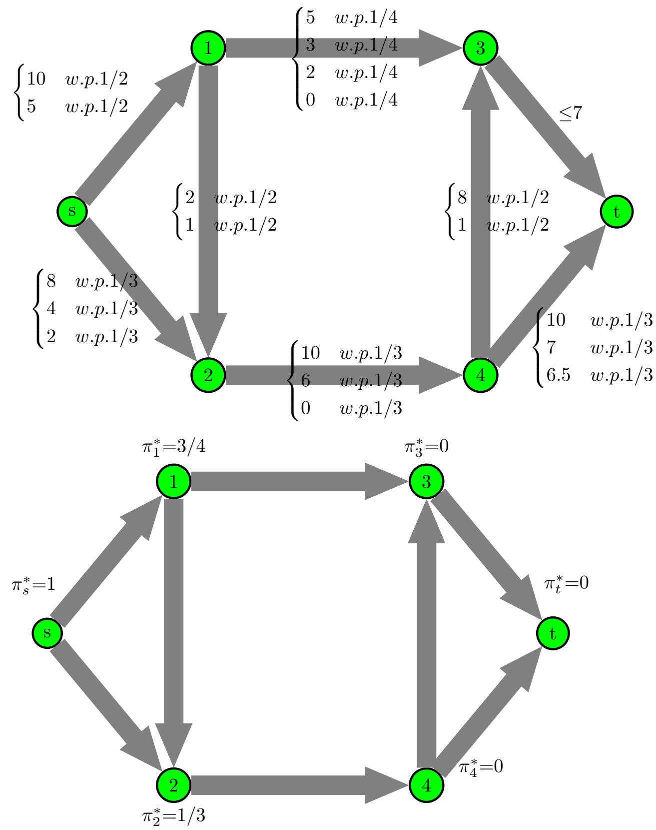

Figure 1: A Distributionally Robust Max Flow instance, and the solution (optimal dual variables π^*) to **Network Dual'**

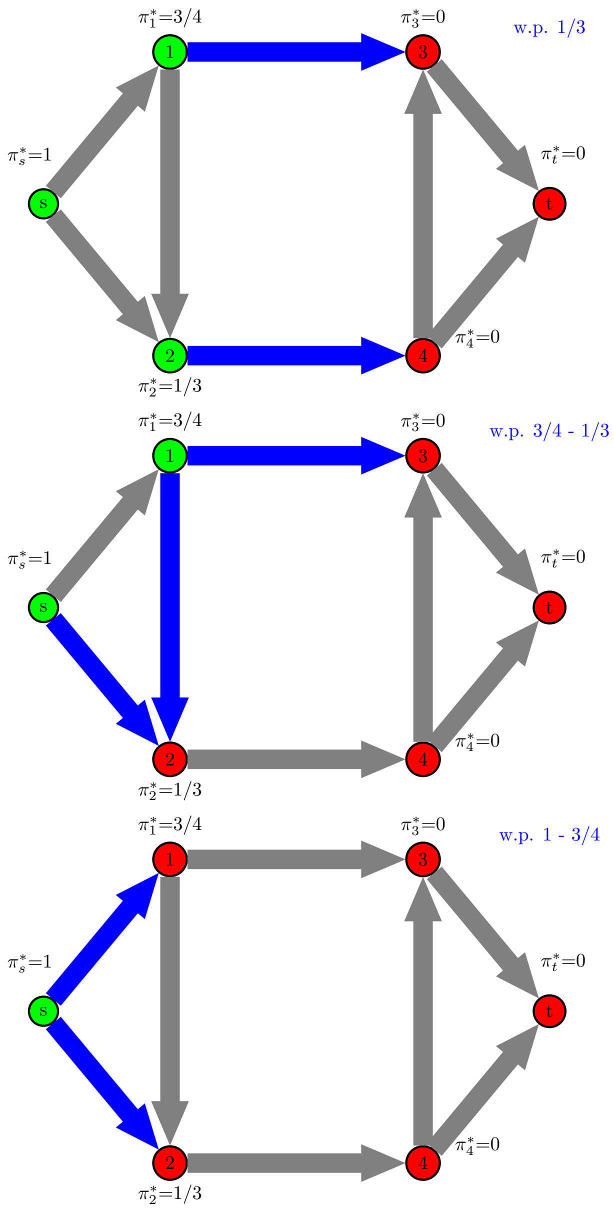

Figure 2: The three graphs depict the distribution over cuts, where $\tilde{S}$ consists of the green nodes, $\tilde{T}$ consists of the red nodes, and the cut-set consists of the blue arcs.

Indeed, if the game were repeated under the stochastic fictitious play dynamic, the cut player's long-run history of selected pure strategies would converge to playing each of the 3 possible s-t cuts with probability 1/3. However, if we allow randomization on the part of the cut player, then we do in fact find that the order of play

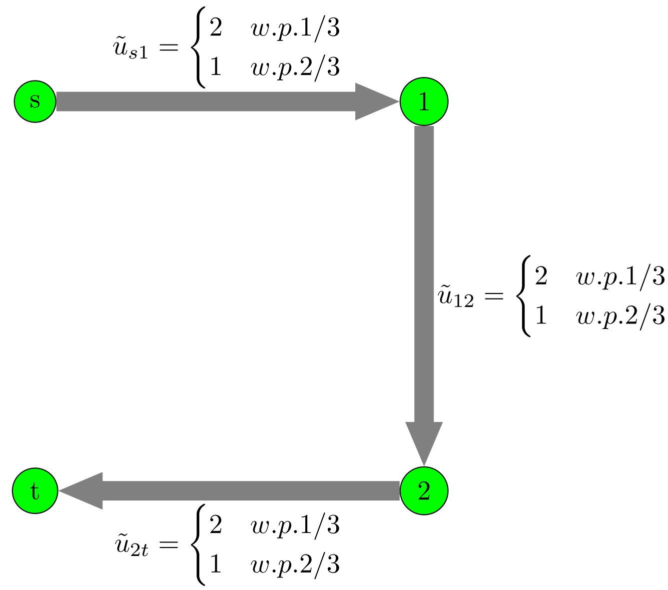

Figure 3: Pure Nash Equilibrium does not exist

does not matter,

$$\begin{aligned} &\max_{\mathbf{w}\, \in [c_0, c_K]^A} \min_{\tilde{\chi}\, \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \mathbb{E}\left[U(\mathbf{w}, \tilde{\chi})\right] \\ &= 1 \\ &= \min_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \max_{\mathbf{w} \in [c_0, c_K]^A} \mathbb{E}\left[U(\mathbf{w}, \tilde{\chi})\right]. \end{aligned}$$

Thus, while a pure Nash equilibrium does not exist here, a mixed Nash Equilibrium does. And this is true in general for this game.

Lemma 5.1. (Existence of Mixed Nash)

$$\max_{\mathbf{w} \in \mathbb{R}^A_{\geq 0}} \min_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \mathbb{E}\left[U(\mathbf{w}, \tilde{\chi})\right] = \min_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \max_{\mathbf{w} \in \mathbb{R}^A_{\geq 0}} \mathbb{E}\left[U(\mathbf{w}, \tilde{\chi})\right]$$

Proof.

$$\begin{aligned} &\max_{\mathbf{w} \in \mathbb{R}^A_{\geq 0}} \min_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \mathbb{E}\left[U(\mathbf{w}, \tilde{\chi})\right] = \max_{\mathbf{w} \in \mathbb{R}^A} \min_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \mathbb{E}\left[U(\mathbf{w}, \tilde{\chi})\right] \\ &= \max_{\mathbf{w} \in \mathbb{R}^A} \min_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \mathbb{E}\left[\mathbf{w}^\intercal \tilde{\chi}\right] - \sum_{(i,j) \in A} \mathbb{E}\left[\max(w_{ij} - \tilde{u}_{ij}, 0)\right] \\ &= \inf_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \max_{w \in \mathbb{R}^A} \mathbb{E}\left[\mathbf{w}^\intercal \tilde{\chi}\right] - \sum_{(i,j) \in A} \mathbb{E}\left[\max(w_{ij} - \tilde{u}_{ij}, 0)\right] \\ &= \inf_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \max_{w \in \mathbb{R}^A_{\geq 0}} \mathbb{E}\left[\mathbf{w}^\intercal \tilde{\chi}\right] - \sum_{(i,j) \in A} \mathbb{E}\left[\max(w_{ij} - \tilde{u}_{ij}, 0)\right] \\ &= \min_{\tilde{\chi} \sim H \in \Delta(\mathcal{X}_{\text{cut}})} \max_{\mathbf{w} \in \mathbb{R}^A_{\geq 0}} \mathbb{E}\left[U(\mathbf{w}, \tilde{\chi})\right], \end{aligned}$$

where the exchange of the "sup" and "inf" to obtain a dual problem is established using a "partial saddle point" result (see Proposition 2.3- Remark 2.3 in [5]).

To conclude the proof, we must argue that the infimum in the dual problem is attained, and this follows because $\max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \mathbb{E}[U(\mathbf{w},\cdot)]$ is convex over $\Delta(\mathcal{X}_{\text{cut}})$, effectively a convex, compact set. ∎

DEFINITION 5.1.
(Capacity Player's Problem)

$$\max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \min_{\tilde{\chi}\sim H\in\Delta(\mathcal{X}_{\text{cut}})} \mathbb{E}\left[U(\mathbf{w},\tilde{\chi})\right]$$

(Cut Player's Problem)

$$\min_{\tilde{\chi}\sim H\in\Delta(\mathcal{X}_{\text{cut}})} \max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \mathbb{E}\left[U(\mathbf{w},\tilde{\chi})\right]$$

Indeed, (1.1) really boils down to analyzing the conflict between the capacity and cut player. We briefly illuminate this connection here.

5.1 The Capacity Player's Problem

COROLLARY 5.1.

$$\begin{aligned}&\max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \min_{\chi\in\mathcal{X}_{\text{cut}}} U(\mathbf{w},\chi)\\&= \max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \Big\{Z(\mathbf{w}) - \sum_{(i,j)\in A} \mathbb{E}_{\tilde{u}_{ij}} \max(w_{ij}-\tilde{u}_{ij},0)\Big\}\end{aligned}$$

Proof. This immediately follows from Lemma 2.1. ∎

5.2 The Cut Player's Problem

COROLLARY 5.2.

$$\begin{aligned}&\min_{\tilde{\chi}\sim H\in\Delta(\mathcal{X}_{\text{cut}})} \max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \mathbb{E}\left[U(\mathbf{w},\tilde{\chi})\right]\\&= \min_{\tilde{\chi}\sim H\in\Delta(\mathcal{X}_{\text{cut}})} \sum_{(i,j)\in A} \int_0^{q_{ij}} F_{ij}^{-1}(p)dp\\&= \min_{\pi\in\Pi_{\text{feas}}} \sum_{(i,j)\in A:\pi_i>\pi_j} \int_0^{\pi_i-\pi_j} F_{ij}^{-1}(p)dp\end{aligned}$$

Proof. Let $\tilde{\chi}\sim H\in\Delta(\mathcal{X}_{\text{cut}})$ be given. Then

$$\begin{aligned}&\max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \mathbb{E}\left[U(\mathbf{w},\tilde{\chi})\right]\\&= \max_{\mathbf{w}\in\mathbb{R}^A_{\geq 0}} \mathbf{w}^{\mathsf{T}}\chi - \sum_{(i,j)\in A} \mathbb{E}\left[\max(w_{ij}-\tilde{u}_{ij},0)\right]\\&= \sum_{(i,j)\in A} \max_{w_{ij}\in\mathbb{R}_{\geq 0}} \Big[w_{ij}\cdot q_{ij} - \mathbb{E}\left[\max(w_{ij}-\tilde{u}_{ij},0)\right]\Big]\\&= \sum_{(i,j)\in A:q_{ij}>0} F_{ij}^{-1}(q_{ij})\cdot q_{ij} - \mathbb{E}\left[\max(F_{ij}^{-1}(q_{ij})-\tilde{u}_{ij},0)\right]\\&= \sum_{(i,j)\in A:q_{ij}>0} \int_0^{q_{ij}} F_{ij}^{-1}(p)dp,\end{aligned}$$

where the equalities follow because it is optimal to set w_{ij} to be $F_{ij}^{-1}(q_{ij})$ when $q_{ij}>0$, and 0 otherwise. This establishes the first equality. The second equality then follows from Lemma 5.1 and Corollary 5.1 ∎

We remark that given any $\pi\in\Pi_{\text{feas}}$, then the $\tilde{\chi}\sim H\in\Delta(\mathcal{X}_{\text{cut}})$ defined via:

$$\tilde{\chi}_{ij}(p) := \begin{cases}1; & \pi_i>\pi_j, p\in[\pi_j,\pi_i]\\0; & \text{otherwise,}\end{cases}$$

for all $(i,j)\in A$, has as its support a sink-side nested (and source-side nested) collection of cut-sets. In other words, there exists a bijection between Π_{feas} and the collection of sink-side nested (and source-side nested) cut-sets. Hence, in light of this, Corollary 5.2 indicates that we (or equivalently, the cut player) do not lose optimality by restricting the search for $\tilde{\chi}$ to those with such nested supports. In the case of finite supported marginals, this search can be conducted efficiently with the help of a linear program- see Section 4.1. This is crucially important for a tractable search for an extremal distribution. Indeed, how to solve the cut player's problem appears, at least at first glance, daunting, considering the pure (much less, mixed) strategy space comprised of all *s*-*t* cut-sets is complex- the problem of merely counting all *s*-*t* cuts is #P Hard ([9]).

References

[1] S. Agrawal, Y. Ding, A. Saberi, and Yinyu Ye. Price of correlations in stochastic optimization. *Operations Research*, 60(1):150–162, 2012.

[2] Shipra Agrawal, Yichuan Ding, Amin Saberi, and Yinyu Ye. Correlation robust stochastic optimization. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1087–1096. Society for Industrial and Applied Mathematics, 2010.

[3] Ravindra K Ahuja, Thomas L Magnanti, James B Orlin, and K Weihe. Network flows: theory, algorithms and applications. *ZOR-Methods and Models of Operations Research*, 41(3):252–254, 1995.

[4] Louis Chen, Will Ma, Karthik Natarajan, David Simchi-Levi, and Zhenzhen Yan. Distributionally robust linear and discrete optimization with marginals. 2018.

[5] I. Ekeland and R. Temam. *Convex analysis and variational problems.* SIAM Classics in Applied Mathematics, Philadelphia, 1999.

[6] P. M. Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, pages 1–52, 2017.

[7] Willem K. Haneveld. *Robustness against dependence in PERT: An application of duality and distributions with*

known marginals. Springer Berlin Heidelberg, Berlin, Heidelberg, 1986.

[8] Karthik Natarajan, Miao Song, and Chung-Piaw Teo. Persistency model and its applications in choice modeling. *Management Science*, 55(3):453–469, 2009.

[9] J Scott Provan and Michael O Ball. The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM Journal on Computing*, 12(4):777–788, 1983.

[10] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

[11] G. Weiss. Stochastic bounds on distributions of optimal value functions with applications to pert, network flows and reliability. *Operations Research*, 34(4):595–605, 1986.

[12] Qiqi Yan. Mechanism design via correlation gap. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 710–719. Society for Industrial and Applied Mathematics, 2011.